Frontline Learning Research
An official journal of EARLI

Frontline Learning Research Vol. 6 No 2 (2018) 20 - 38
ISSN 2295-3159

'A double-edged sword. This is powerful but it could be used destructively': Perspectives of early career education researchers on learning analytics

Jenna Mittelmeier[a], Rebecca L. Edwards[b], Sarah K. Davis[b], Quan Nguyen[a], Victoria L. Murphy[a], Leonie Brummer[c], Bart Rienties[a]

[a]The Open University, UK
[b]University of Victoria, Canada
[c]University of Groningen, The Netherlands

Article received 2 March 2018 / Revised 8 June / Accepted 3 August / available online 17 August

Abstract

Learning analytics has been increasingly outlined as a powerful tool for measuring, analysing, and predicting learning experiences and behaviours. The rising use of learning analytics means that many educational researchers now require new ranges of technical analytical skills to contribute to an increasingly data-heavy field. However, it has been argued that educational data scientists are a 'scarce breed' (Buckingham Shum et al., 2013) and that more resources are needed to support the next generation of early career researchers in the education field. At the same time, little is known about how early career education researchers feel towards learning analytics and whether it is important to their current and future research practices. Using a thematic analysis of a participatory learning analytics workshop discussions with 25 early career education researchers, we outline in this article their ambitions, challenges and anxieties towards learning analytics. In doing so, we have provided a roadmap for how the learning analytics field might evolve and practical implications for supporting early career researchers' development.

Keywords: Learning analytics, educational research, early career researchers

Corresponding author: Dr Jenna Mittelmeier, Institute of Educational Technology, Walton Hall, Open University, Milton Keynes MK5 8EY, United Kingdom. jenna.mittelmeier@open.ac.uk. DOI: 10.14786/flr.v6i2.348

1. Introduction

There is increased awareness in education that using learning analytics to measure students' and teachers' attitudes, behaviours, and cognition provides potential opportunities to enhance and enrich learning and teaching (Agudo-Peregrina, Iglesias-Pradas, Conde-González, & Hernández-García, 2014; Dawson & Siemens, 2014; Tempelaar, Rienties, & Giesbers, 2015; van Leeuwen, Janssen, Erkens, & Brekelmans, 2015; Winne, 2017). Furthermore, fine-grained learning analytics data can provide opportunities to test, validate, or build educational theories (Malmberg, Järvelä, & Järvenoja, 2017; Winne, 2017). Despite these potentials, large-scale uptake of learning analytics has so far been rather limited (Ferguson et al., 2016; Papamitsiou & Economides, 2016). This is reflective of wider scepticism about whether learning analytics-supported teaching and, more generally, technology-enhanced learning, can be as effective as 'traditional' methods (Kirschner & Erkens, 2013; Selwyn, 2015, 2016).

In this regard, a group of prominent researchers in learning analytics and educational data mining recently argued that educational data scientists were a 'scarce breed'(Buckingham Shum et al., 2013). They further noted that critical reflection was needed on how the educational research community could train, support, and prepare the next generation of early career education researchers to work across traditional and new boundaries of research (Buckingham Shum et al., 2013). However, in light of this suggestion, questions remain around how early career education researchers reflect and act upon an increasingly complex, data-heavy research environment. To answer this, we conducted an in-depth, participatory workshop with 25 early career education researchers at the EARLI JURE 2017 conference, which collaboratively and critically evaluated the affordances and limitations of learning analytics. In particular, we outline through a thematic analysis of workshop discussions the ambitions, challenges, and anxieties of early career researchers towards learning analytics, and how and why they might (not) use learning analytics methods in their own research now and in the future.

2. Learning analytics: six main research topics

Learning analytics can be defined as 'the measurement, collection, analysis, and reporting of data about learners and their contexts for the purpose of understanding and optimising learning and the environments in which it occurs' (LAK, 2011). Recent literature in the learning analytics field addresses a large range of trends and challenges in this emerging research field (Ferguson et al., 2016; Lang, Siemens, Wise, & Gašević, 2017; Papamitsiou & Economides, 2016; Selwyn, 2015; Siemens, 2013). For example, in the recent Handbook of Learning Analytics, Lang et al. (2017) identified 27 topics of interest. Given the limited time available for the workshop, we aimed to narrow this work to outline overarching topics of importance that represented cutting-edge debates in the learning analytics field. To aid in this process, we cross-referenced this list with two recent meta-analyses of learning analytics (Ferguson et al., 2016; Papamitsiou & Economides, 2016), which have provided in-depth reviews of current findings, debates in the field, and limitations of recent work. To select topics that were timely to the field, we considered common themes presented at the recent Learning Analytics and Knowledge conferences (LAK), which all workshop facilitators had previously attended. Altogether, these various sources of evidence and expertise were combined to narrow down and identify six main research topics in learning analytics to focus upon in this workshop, which are outlined in Table 1.

Table 1
Six topics of discussion during EARLI JURE learning analytics workshop

Topic Number	Topic
Topic 1	Learning analytics measurements (Agudo-Peregrina et al., 2014; Tempelaar et al., 2015)
Topic 2	Ethics (Siemens, 2013; Slade & Prinsloo, 2013)
Topic 3	Learning design (Lockyer, Heathcote, & Dawson, 2013; Nguyen, Rienties, Toetenel, Ferguson, & Whitelock, 2017; Rienties & Toetenel, 2016)
Topic 4	Learning dispositions (Buckingham Shum & Deakin Crick, 2012; Tempelaar, Rienties, Mittelmeier, & Nguyen, 2017)
Topic 5	Combining learning analytics with other methods (Chatti, Dyckhoff, Schroeder, & Thüs, 2012; Malmberg et al., 2017)
Topic 6	Generalisation versus personalisation (Ifenthaler & Widanapathirana, 2014; Scholes, 2016)

In the next sections, a description and justification for using these six topics in our workshop are provided, followed by a description of our research questions for this study.

2.1 Learning analytics measurements

The opportunity to capture and measure rich and fine-grained data from authentic learning environments in real-time has been one of the unique advantages of learning analytics research. For example, in a large-scale study amongst 141 courses, Rienties and Toetenel (2016) found that the way teachers designed courses significantly predicted how 111,000 students were studying online, whereby follow-up training with 80+ teachers helped to improve the learning design. However, since the data are often collected authentically (versus in controlled lab settings) and automatically (as opposed to collection first-hand by researchers), questions about how to compute meaningful metrics remain. For example, the typical log files that record user activities on a browser lack insight into many factors, such as who is behind the screen, whether they are active or taking a break, or whether they are focused on task or distracted (Kovanovic, Gašević, Dawson, Joksimovic, & Baker, 2016; Kovanovic et al., 2015). Without explicit engagement with the underlying assumptions behind these data, researchers could overestimate or underestimate results (Gašević, Dawson, & Siemens, 2015). Therefore, it is crucial for early career education researchers to be aware of both the strengths and limitations of these measurements to adopt learning analytics in meaningful manners.

2.2 Ethics

Several ethical debates have emerged around the practical application of learning analytics techniques. For example, questions remain around transparency of how and why data is collected and used (Slade & Prinsloo, 2013), what information is missing from automatically captured data (Ruppert et al., 2015), what biases and subjectivities are inherent to techniques used for analysis (boyd & Crawford, 2012), and whether educational institutions have an obligation to utilise learning analytics to support students (Prinsloo & Slade, 2017a). Perhaps most prominently, there are criticisms around data ownership and students' consent for their data to be collected. Whether data are owned by universities, students, or even researchers has large implications for the ethicality of learning analytics (Drachsler & Greller, 2016). Furthermore, generic consent is often collected at enrolment, but, as a wide variety of research projects may potentially use students' data in new and various ways, it is difficult to argue that this is *informed* consent (Slade & Prinsloo, 2013). As the field struggles to answer these questions, it is critical that early career education researchers engage in the debates surrounding ethical learning analytics.

2.3 Learning design

In recent years, there has been increasing interest in the connection between learning design and learning analytics. In his LAK16 keynote speech, Kirschner (2016) critically reflected on the lack of pedagogical context in learning analytics research, which could prevent researchers from asking the right questions, selecting the right metrics, or interpreting results in meaningful manners (Gašević et al., 2015). In this way, learning design is an emerging field that aims to develop a 'descriptive framework for teaching and learning activities ("educational notation"), and to explore how this framework can assist educators to share and adopt great teaching ideas' (Dalziel, 2016). By capturing and visualising the design of learning activities, learning design could provide pedagogical context to support interpreting and translating learning analytics findings into tangible interventions (Lockyer et al., 2013; Persico & Pozzi, 2015). A gradual accumulation of empirical evidence has reinforced the importance of aligning learning analytics with learning design, as the way teachers design for learning has been found to be significantly related to students' engagement, pass rates, and satisfaction (Nguyen et al., 2017; Rienties & Toetenel, 2016). In this way, learning design offers opportunities for early career education researchers to engage with educational theories and pedagogies that underpin their analysis.

2.4 Learning dispositions

Early learning analytics research often focused on predictive models based on extracting data from digital platforms (such as learning management systems) and institutional records of student information. While these studies (e.g., Agudo-Peregrina et al., 2014) provided valuable foundations for understanding learning analytics' potential, they relied primarily on demographics, grades, and trace data, providing a relatively simplistic story of students' experiences and behaviours. A reliance on such narratives without more complex insights may lead to difficulties in designing pedagogically-informed interventions (Tempelaar et al., 2015; Tempelaar et al., 2017). In response, Buckingham Shum and Deakin Crick (2012) proposed a 'dispositional learning analytics' approach that combines behavioural data (e.g., that which is mined from learning management systems) with more complex data about the learner (e.g., their values, attitudes, and dispositions measured through self-reported surveys) (Malmberg et al., 2017; Tempelaar et al., 2015). For example, recent research in blended courses of mathematics and statistics (Tempelaar et al., 2017) found that linking learning dispositions data with behavioural learning analytics data can lead to actionable feedback. As the learning analytics field matures, therefore, it becomes more necessary for early career education researchers to understand and apply more complex learner data for use in tandem with learning analytics.

2.5 Combining learning analytics with other methods

In light of the issues outlined above, combining learning analytics with other methods offers opportunities for both early career researchers and the wider field. For example, it has been argued that combining more classically automated and quantitative learning analytics data with qualitative methods (interviews, focus groups, think-aloud protocols, etc.) can develop insight into the *how* and *why* factors impacting quantitative findings (Chatti et al., 2012). Similarly, Merceron, Blikstein, and Siemens (2016) described learning as 'multimodal' and involving many simultaneous mental and physical processes, noting maturity in learning analytics research that incorporates multiple methods to capture this. Some potential additional methods include natural language processing, eye tracking, social network analysis (such as through surveys), visualisation techniques, discourse analysis, or emotional measurements (Suthers & Verbert, 2013). Combining learning analytics with these techniques reflects the diversity of the learning analytics field as a whole, which crosses multiple academic disciplines (Dawson, Gasevic, Siemens, & Joksimovic, 2014). This diversification of learning analytics research has implications for early career education researchers as they develop skills and competencies to contribute to the field.

2.6 Personalisation versus generalisation

Learning analytics applications can be continuously adapted and, to some extent, personalise the learning environment for students (Ifenthaler & Widanapathirana, 2014). However, personalisation relies on categorising students based on behaviours and traits (Buckingham Shum & Deakin Crick, 2012), and some have noted that there are grey areas between 'categorising' and 'stereotyping' (Prinsloo & Slade, 2017b; Scholes, 2016; Slade & Prinsloo, 2013). For instance, learning analytics could lead to erroneous assumptions about future behaviours and performance (Scholes, 2016). Issues also remain around the bias of researchers or teachers, whose views towards certain students or categories of students might influence the interpretation of findings. Therefore, it is important for early career education researchers to be explicitly cognisant of the fine line between personalisation and generalisation to develop sustainable algorithms and interpretations that limit researcher bias.

3. Method

3.1 Purpose and Research Questions

Current discussions in learning analytics research, as outlined by the six topics above, depict a complex and diverse emerging field. As such, the next generation of early career education researchers represent an important voice in understanding how learning analytics research will develop and progress (Buckingham Shum et al., 2013). In this study, we aimed to understand how these early career researchers make sense of learning analytics in light of these ongoing debates and discussions in the field. Using an interactive workshop structured around these six learning analytics topics amongst 25 early career education researchers in the education field, we explored the following research questions:

1. What are the main strengths and limitations of learning analytics according to early education career researchers?
2. In what ways do early career education researchers embed learning analytics approaches into their own research practices?

In answering these questions, we have outlined an in-depth account of how early career education researchers approach learning analytics techniques, providing insight into potential pathways forward for the field.

3.2 Procedure and Setting

This study describes the results of an invited workshop at the EARLI JURE 2017 conference in Tampere, Finland, which was entitled 'Three different perspectives on why you need learning analytics and educational data-mining'. EARLI (European Association for Research on Learning and Instruction) is an international organisation for researchers in the education field, and JURE (Junior Researchers of EARLI) focusses specifically on early career researchers in education. Early career researcher in this context is defined as Master's students, PhD students, and those within two years of receiving their PhD. More information about the conference and its aims are available at: http://www.earli-jure2017.org/

The workshop was facilitated by three researchers from The Open University (OU), including a PhD student, a postdoctoral researcher, and a professor. The workshop was 90 minutes long and consisted of opening and closing sessions, combined with breakaway small group discussions for each of the six topics outlined in the introductory sections of this article (see Table 1). Participants were free to choose which small group topic they wished to discuss and not all participants discussed all topics. On average, each participant attended two small group discussions. The overall aim of the workshop was to encourage critical assessment and debate on the merits and drawbacks of learning analytics approaches in participants' own institutional and cultural contexts. Table 2 describes the workshop schedule adopted.

Table 2
Workshop schedule outline

	Activity	**Duration**	**Description**
1	Introduction to learning analytics and measurement of initial expertise/expectations	5 minutes	Introduction of aims of the workshop, defining learning analytics, and an interactive online survey
2	Open discussion on strength and weakness of learning analytics	15 minutes	Open large group discussion on strengths and limitations of learning analytics
3	Small group discussions	60 minutes	Small group discussions around the six topics, during which time participants freely moved to other topics as interested. During this time, discussions on three topics were held concurrently and participants were free to choose a topic that most interested them. After 30 minutes, the three topics under discussion changed and participants could once again choose a new discussion topic according to their interests.
4	Bringing together perspectives	10 minutes	Bringing together discussions and perspectives of participants on the six topics

The workshop was highly participatory with opportunities for attendees to take part in both small and large group discussions. A pilot was conducted two weeks before the EARLI JURE workshop with 12 early career researchers at The Open University to test and fine-tune the workshop design, as well as the final selection of the six topics. The role of the facilitators was mostly limited to moderating and encouraging discussion, as the aim was to offer collaborative experiences between early career education researchers. The workshop facilitators only provided a one-minute introduction of each of the six topics outlined in Table 1 before breaking into small groups and contributed to discussions only to facilitate conversation between participants (i.e., 'What are your thoughts on this topic?' or 'How does this relate to your own work?'). The full slides used by the facilitators during the workshop is available online (see: https://bit.ly/2Mep4U0).

This design allowed early career researchers to contribute their own voices and experiences, as well as reflect upon their own practices. In terms of the six learning analytics topics (Table 1), our aim was to encourage discussion around current debates on these topics while creating an environment that allowed participants to be critical about their relevance and contribute their own opinions. The workshop design further allowed participants to draw upon their own needs as researchers, including what their practices had in common with learning analytics in the wider field, as well as what is unique about their own experiences. During the full group discussion at the start of the workshop (Activity 2 in Table 2), 19 of the 25 participants contributed their opinion about the strengths and limitations of learning analytics. In the small group discussions (Activity 3 in Table 2), each of the 25 participants made at least one comment about their chosen small group discussion topic. Altogether, all participants were active during the workshop and most provided substantial discourse about their opinions and experiences.

3.3 Participants

Workshop participants were recruited via the EARLI JURE conference registration process and 25 conference attendees joined the session. Overall, there was a wide range of participants, including PhD students, postdoctoral researchers, those from non-academic sectors (e.g., non-profit or regulatory bodies), and policymakers. In terms of geographical spread, participants were from Belgium, Canada, China, Germany, Finland, the Netherlands, Pakistan, Spain, Singapore, Vietnam, the United Kingdom, and the United States. Although all participants were conducting education-related research, they also came from a variety of disciplines (e.g., educational technology, computer science, learning sciences, primary education, secondary education, higher education, educational psychology, etc.). Following the workshop, four participants collaborated with the research team as co-researchers in this study and contributed to the analysis and writing. The research team was a diverse mix of early career researchers from Canada, The Netherlands, UK, USA, and Vietnam, who conducted research in Canada, the Netherlands, and the UK. This collaboration ensured the credibility and authenticity of the findings and contributed to the validity by 'building the participant's view into the study' (Creswell & Miller, 2000, p. 128).

3.4 Instruments

3.4.1 Online surveys

An online survey tool (pollev.com) was used to probe six questions to participants during the workshop. These questions are outlined in Table 3. In total, 17 to 20 participants completed each of the questions at the start of the workshop (questions 1-4), 22 participants completed the topic ranking activity (question 5), and 13 participants completed end-of-workshop satisfaction questions (questions 6-7). Descriptive statistics of our findings from these questions are compiled in the results section to provide a general understanding of participants' backgrounds and views on learning analytics and its applicability to their own research.

Table 3
Survey questions asked of participants

Number	Timing	Question
1	Start of workshop	What is the first word that pops up into your mind when you hear about learning analytics? (short answer)
2	Start of workshop	How well do you understand what learning analytics means and involves? (1-5 Likert scale)
3	Start of workshop	How relevant is learning analytics to your research? (1-5 Likert scale)
4	Start of workshop	How important do you feel it is for universities to collect and analyse learning analytics data about their students? (1-5 Likert scale)
5	Prior to small group discussions	Which of the six research themes would you most like to discuss? (ranking of six topics)
6	End of workshop	How satisfied are you with the workshop's content? (1-5 Likert scale)
7	End of workshop	How satisfied were you with the workshop's approach? (1-5 Likert scale)

3.4.2 Transcripts of discussions

Both large and small group discussions were recorded during the workshop and transcribed verbatim. The workshop facilitators also wrote discussion points and perspectives from participants on flipcharts, which served as a secondary data source for qualitative analysis. In total, eight transcripts and three documents were used to inform our findings. Although workshop participants were split into small groups to discuss different topics during Activity 3 (Table 2), we were interested in understanding the common and overarching themes that emerged across the six discussion topics. As such, we used a thematic analysis method, as described by Lichtman (2013), to summarise and group common themes across all the data available from this workshop, including whole group discussions, small group discussions, workshop notes, and facilitator notes and reflections. For the first stage of the thematic analysis, each member of the research team provided an initial summary of key emergent themes for one of the transcripts or data sources, which were then each reviewed and confirmed by two additional members of the research team for accuracy and reliability. In the next stage, the themes from each data source were then combined to develop an overarching coding book of common themes across all data sources (as outlined in Appendix 1). This two-stage analysis provided an understanding of participants' sentiments both within and across the whole and small group discussions. Afterwards, notes and reflections on these themes were compared between all authors to confirm and validate findings, counter any disagreements, and further develop the narrative of our findings.

3.4.3 Post-workshop reflections

Directly after the workshop, the three workshop facilitators and a selection of workshop participants met to reflect upon the discussions and the main themes that emerged within and across the six discussion topics. This conversation was recorded and transcribed to inform the data analysis described in Section 3.3.2.

4. Results

The descriptive statistics from the online survey provided an overview of participants' backgrounds and views towards learning analytics. The first question asked on a 1-5 scale the degree of understanding participants felt they had about learning analytics (1 = no idea and 5 = expert). To this, only 15% (n = 3) considered themselves experts, while the majority (80%, n = 16) rated their understanding was somewhere in the middle (between scales 2 – 4). We next asked whether participants felt that learning analytics was relevant to their own research on a 1-5 scale (1 = irrelevant, 5 = extremely relevant) and the majority (80%, n = 16) agreed (scored 3 and above). Most participants (83%, n = 17) also felt that learning analytics are of importance for their universities (scored 3 and above). Overall, despite a large variation in terms of the level of familiarity with learning analytics, the majority of early career education researchers indicated that learning analytics is relevant to their work and institutions.

4.1 RQ 1. Main strengths and limitations of learning analytics

In the opening large-group discussion, participants outlined a number of key strengths and weaknesses of learning analytics, which is summarised in Table 4. Perspectives were provided from 19 of the 25 participants during this portion of the workshop, indicating high participation across the full sample.

Table 4
Main strengths and limitations from the opening session at the JURE workshop

Strengths	(1) Harnessing the power of data (2) Learning analytics as 'objective' (3) Ability to measure multiple things simultaneously (4) Ability to personalise feedback to learners (5) Opportunities for cross-institutional collaboration to test/validate educational theories
Limitations	(1) Learning analytics presenting an incomplete picture of learning (2) Strong technical and statistical skills required

(3) Cognitive overload for teachers and students trying to understand/use findings
(4) Ethical use of data and restrictions on data usage
(5) Dangers of categorising students inappropriately

In terms of strengths, participants indicated first and foremost that learning analytics data can provide insights into actual learner behaviours. Six whole-group discussion participants, in that way, argued that learning analytics data is more 'objective.' For example, one participant commented, 'it's not relying on self-report.' Of course, whether learning analytics data is actually objective can be debated (Mirriahi & Vigentini, 2017; Prinsloo & Slade, 2017b). This point was further articulated by one participant, who suggested '*perceived* objectivity…because you have to do something with the data to get something out of it.' In addition, the power to visualise data and provide immediate feedback to both teachers and students was frequently noted by several participants, as highlighted by recent research (Charleer, Klerkx, Duval, De Laet, & Verbert, 2016; Nguyen et al., 2017). Furthermore, two additional participants noted that learning analytics can measure 'a lot of things at the same time,' while most educational researchers primarily have to be selective in terms of the instruments and constructs that can be collected or measured. Another potential strength highlighted by two participants was the opportunity to give direct, personalised feedback to learners, or so-called formative learning analytics (Sharples et al., 2016). Finally, one participant noted that learning analytics research may provide cross-institutional collaboration opportunities to test and validate large and small educational theories (Winne, 2017). Altogether, it was apparent that early career participants were engaged with the potential power of learning analytics and felt that it had value for the future of educational research, as indicated by over half of the comments provided during the whole group discussion.

In terms of limitations (also summarised in Table 4), there were fears from four participants that parameter-driven analytics might lead to 'false data', which potentially could lead to wrong conclusions or inappropriate interventions based on an incomplete picture of learning. Additionally, concerns were voiced by five participants around the potential for learning analytics to seem like a 'black box' (García et al., 2012; Kovanovic et al., 2015). Four participants mentioned that learning analytics required an intimidating level of statistical techniques and technical competencies, which had implications for their skills development pathways and overall perception of the field amongst peers. Relatedly, another comment outlined potential risks of cognitive overload for teachers and students when making sense of data. Ethics was also a frequent cause for concern. In particular, four participants were worried about 'who owns the data' - students, teachers, institutions, governments, companies? There were some additional concerns around informed consent and whether learners were aware that their data was being collected (Slade & Prinsloo, 2013). At the same time, they noted difficulties in that some countries or institutions limited or even forbade measuring and monitoring students' and teachers' activities (Prinsloo & Slade, 2017b; Slade & Prinsloo, 2013), which was detrimental to their research. A final but crucial limitation outlined by the final comment that links with all points above was the risk of labelling (i.e., 'the oversimplification of student characteristics'), whereby students (or teachers) may be identified to fit in a particular category of users based upon a mix of objective and subjective indicators and algorithms. Altogether, nearly half of the discussion points made by early career participants in the whole-group discussion appeared to be engaged with the current debates in the field and felt that creative solutions were needed to both accomplish their personal research goals and drive the field forward.

4.2. RQ 2. Embedding learning analytics into research practice

In the online survey, we asked workshop participants to rank their personal interest in the six identified learning analytics topics. The subjects were collectively ranked in the following order: (1) learning analytics measurement, (2) learning design, (3) combining learning analytics with other methods, (4) generalisation versus personalisation, (5) ethics, and (6) learning dispositions. This seemed to initially indicate that early career education researchers were more strongly engaged with issues pertaining to methods and methodologies over those of more theoretical importance. One explanation could be that many of the students were PhD students who were interested in discussing and framing their research projects and methodologies.

In the small groups, discussions centred on the chosen topic with a focus on how participants were currently integrating learning analytics into their own research practices. Potential future uses of learning analytics to enhance the participants' research projects were also explored in the dialogue in light of the small group topic. Although discussions varied across the small groups and it was not possible to elicit whether sentiments were shared across all participants, our thematic analysis identified several common themes across all six small group discussions in this workshop. Altogether, the majority of early career education researchers in this workshop:

1. Used learning analytics to support student success through self-regulated learning approaches
2. Felt that learning analytics were on the cutting edge of educational research, which posed unique challenges for their own research and the wider field
3. Believed that approaches of combining learning analytics data with other methods strengthened learning analytics and aimed to incorporate multiple methods into their own research
4. Contemplated theory-driven learning analytics versus learning analytics-driven theory
5. Aimed to contextualise their data to the realities of students' lives and experiences
6. Recognised a multitude of ethical ramifications for their own learning analytics research

Definitions of each theme and its subthemes are provided in Appendix 1. The following sections provide a detailed narrative of each theme derived from our analysis.

4.2.1 Theme 1: Using learning analytics to support student success by taking a self-regulated learning approach

'There is a lot of use of demographic data in LA and I like what Phil Winne said in the Handbook of Learning Analytics…that he thinks the focus should be on things learners can change about themselves.'

In all six of the small discussion topic groups, participants suggested that combining learning and learner data provides opportunities to support student success, both for students who are at risk of failure and students who are objectively successful but could achieve more. In four of the six small groups, this discussion was framed through the perspective of researching things that students have agency over (e.g., behaviours that they can change) rather than static demographic variables. By focusing on these types of variables, several participants noted that it might be possible to provide just-in-time and just-enough feedback to students so that they can strategically adjust their learning approaches. Particularly in Topic 4, 5, and 6 discussions, participants suggested that it is critical for students to take an active role and learn to effectively regulate their own learning, thereby not use learning analytics as a 'crutch.' In this way, participants in these groups argued that it is important to 'sell' learning analytics as beneficial to students, not just advisors and teachers.

4.2.2 Theme 2: Learning analytics are on the cutting edge of educational research and this poses unique challenges

'For right now, it's just me and I have to find everyone. So, I'm building it from the ground up and I have to find people at my institution with these skills.'

In the whole group discussion and our workshop survey, participants noted there are wide variations in levels of familiarity and comfort with learning analytics techniques at their institutions. In five of the six small groups, participants elaborated that, despite this inconsistency, they personally viewed learning analytics as a powerful method because it brings together researchers from diverse disciplines in a way that traditional educational research methods have not. However, participants have found it challenging to establish (or find) needed expertise, work with complex data, learn and use advanced analysis techniques, and access needed technologies. This notion was highlighted in the whole-group discussion, the workshop survey, and in all six of the small group discussions. Across the various discussions, there was a sense that these challenges stemmed from the newness of the field and participants often likened using learning analytics to embarking alone on an adventure for which you are not prepared.

4.2.3 Theme 3: Combining learning analytics methods with other methods

'I ran in a lot of trouble of, like, getting a lot of data but not really feel how to interpret it. Because you only see the patterns, but you don't know why or how … I'm really thinking about mixing it, with the log data analysis but then also interviewing them … You usually need somebody's explanation about why they did what they did.'

Many participants in both the whole and small group discussions consistently noted the value of combining learning analytics with more traditional education research methodologies. Although this was particularly prevalent in Topic 5 (combining learning analytics with other methods) and Topic 4 (learning dispositions), the value of using other methods in combination with the more automated and quantitative learning analytics data was outlined in three other small group discussions. When discussing their own work throughout the workshop, comments from 15 of the 25 participants indicated that early career education researchers were either already multiple methods or hoped to combine methods in the future. Among other methods, participants referenced social network analysis surveys, interviews, document analysis, eye tracking technologies, and psychometric questionnaires. In general, participants aimed to combine learning analytics with other methods because they believed that this would provide a stronger evidence base, mitigate biases, and add the 'how-and-why' or intent to their analysis and findings. However, several participants in the Topic 5 discussion expressed concerns that they may not be able to access the data or participant numbers they would require to use multiple methods.

4.2.4 Theme 4: Contemplating theory-driven learning analytics versus learning analytics-driven theory

'…The theory often seduces you to see something in the data which might not entirely be there. It's still quite tricky…'

There was some debate in the whole group discussions (four comments) and across the small group discussions (twelve comments) on whether learning analytics should be theory-driven or data-driven. Discussed by nearly half of the participants throughout the workshop, there were conflicting views between them about whether theory should inform data analysis choices or whether the data should be explored to develop or inform new theories. Four participants described that focusing on existing theories was limiting, as the data should 'speak for itself'. In this way, there were suggestions that theories should be developed from the analysis and findings because existing theories might not match students' measurable behaviours. At the same time, five other participants gravitated towards more theory-driven analytics to understand what variables to consider and to bring the 'why' factor into their findings. This second perspective was more in line with more established experts in the field (Gašević et al., 2015), who have argued that learning analytics researchers must explicitly engage with the underlying assumptions of using learning analytics data and the educational theories underpinning their research.

4.2.5 Theme 5: Aiming to contextualise their data to the realities of students' lives and experiences

'Sometimes it's not the amount of time they spent, but the sequence of time they spent on something. Or you have to know the time when they have an exam, so you need to take into account the context of the course.'

Throughout all workshop activities, participants aimed to contextualise the data they collected in light of students' realities and experiences. This was an explicit focus in the Topic 6 small group discussion, but these sentiments were brought up in the remaining five small groups as well. In all groups, participants suggested that data is easier to understand in context and that findings do not always translate between contexts. This reflects positively on the future of the field, as nearly all participants recognised a need for complex, wellrounded research that goes beyond simple click data to capture the nuances of student experiences. At the same time, several participants, particularly in the Topic 2, 3, and 5 discussions, noted limitations in their own access to data to accomplish this, which ultimately may skew results. For example, two participants commented that data may only be available from learning management systems, while students are also using another platform for their learning. Early career education researchers may also not have access to participants for interviews or other studies. Therefore, the ability to access data from multiple platforms and gain access to participants were key areas of concern for several participants.

4.2.6 Theme 6: Recognising a multitude of ethical ramifications

'It's about the argumentation of why you focus on the women rather than the men, maybe. For example, if the beneficial effects are larger for women than for men, then I would say, yeah, maybe. But, then again....'

Across the whole group and all six small group discussions, most participants were keenly aware that strong ethics policies are needed when using learning analytics. Although a key focus in the Topic 2 discussion, participants in all small groups were concerned with protecting privacy and individual rights and frequently suggested that it might not be ethical to track everything that one might be capable of tracking. In the Topic 6 discussion, participants indicated that learning analytics, particularly sharing learning analytics data back to teachers and students, might lead to unintended consequences. In the whole group discussion, one participant went so far as to call learning analytics a 'double-edged sword,' pointing out that learn analytics are 'powerful, but could be used destructively'. In this way, participants throughout the workshop argued that learners, educators, researchers, and institutions might not be fully prepared to wield learning analytics in a useful and productive way. However, several participants, particularly those in Topic 1, 2 and 5, did point out that the use of learning analytics has ethical *benefits*. For example, these participants felt that learning analytics provides a way to track and support all students to *reduce* bias and inequalities. Altogether, There was an engagement across the workshop by most participants with ethical implications of learning analytics in line with current debates in the field (Gašević, Dawson, & Jovanovic, 2016; Prinsloo & Slade, 2017b; Slade & Prinsloo, 2013).

5. Discussion and Practical Implications

Previously, researchers have outlined a need for preparing early career education researchers for an increasingly data-heavy education sector (Buckingham Shum et al., 2013). In response, our thematic analysis of the EARLI JURE learning analytics workshop activities have provided a detailed and nuanced account of early career education researchers' feelings towards learning analytics (RQ1) and the role it plays in their own research practices (RQ2). Our findings overall indicated that the early career participants were engaged with current debates and discussions in the field, were interested in developing creative solutions to overcome perceived issues and desired support for building their skills and expertise on the topic.

Workshop participants were broadly in agreement about several themes, including the potential of learning analytics to support learning, the value of combining learning analytics with other methods, the need to contextualise learning analytics to learners' own lives, and the need to engage with ethical implications. These views are in line with experts in the field and indicate that most early career education researchers using learning analytics are in tune with the wider community of researchers (Ferguson et al., 2016; Papamitsiou & Economides, 2016). At the same time, there were inconsistencies between participants in their views around issues such as the connection between theory and data and whether quantitative learning analytics data is 'objective', which was often in contrast to recent theoretical developments in the learning analytics field (Dawson & Siemens, 2014; Gašević et al., 2015; Kovanovic et al., 2015; Lockyer et al., 2013). This suggests the need for early career education researchers to engage more with learning analytics theory in addition to its methodological and analysis affordances. The sentiments expressed during the workshop appeared to be consistent across participants and could potentially be generalisable to other emerging learning analytics researchers, particularly considering the wide diversity of researchers in attendance. However, it is important to note that most participants were already using or considering using learning analytics, and more diverse views or scepticism would likely be present among those who do not use or have more limited knowledge of learning analytics.

Drawing on the findings we have described, this study reveals four important areas for supporting early career education researchers in the learning analytics field: 1) providing adequate statistical and technical training, 2) building on their curiosities about the bigger picture of learning analytics research, 3) supporting access to different kinds of data, and 4) developing and fostering engagement with ethical issues. Using a combination of large and small group interactions in this study, we revealed that early career researchers echo current debates frequently discussed by learning analytics experts, but are also concerned about how they can contribute their own perspectives to this emerging field.

Providing adequate statistical and technical training

Learning analytics research requires a certain amount of statistical and computational understanding (Lang et al., 2017) and current training for education researchers does not always adequately prepare them to conduct these analyses confidently. If early career education researchers want to learn the skills needed to conduct learning analytics research, they often have to pursue this on their own time with little direction or guidance on what they should be learning. This was made evident in our study, whereby early career education researchers modestly rated their knowledge of learning analytics and expressed concerns and anxieties related to their own statistical and technical skills. The comments also link with recent work, which has highlighted poor data literacy as a key barrier to learning analytics adoption (Ferguson et al., 2014). Therefore, resources are needed to support early career education researchers in building capacities to interpret, analyse, and evaluate big (and small) data. Collaborating on learning research projects with other experienced learning analytics researchers could also help demystify the processes used and provide opportunities for contributions from early career researchers.

Building on early career researchers' curiosities about the bigger picture of learning analytics research

In this study, early career education researchers showed concerns about learning analytics research presenting an incomplete picture of students' learning. Specifically, they were concerned about the oversimplification of students' characteristics and using demographic data collected about students to inform interventions. Participants recognised learning analytics might provide only a snapshot of students' behaviour and performance. Additionally, they wondered how data collected could be used to inform theories of learning (and vice versa) and how variables represent different aspects of learning. These echo common concerns and theoretical arguments that are being developed by experts in the wider learning analytics field (Dawson & Siemens, 2014; Gašević et al., 2015; Kovanovic et al., 2015; Lockyer et al., 2013). As such, PhD and postdoctoral research projects represent a valuable resource for pushing the boundaries of the learning analytics field. Supervisors and line managers should, therefore, build on the curiosities of early career education researchers by supporting theoretical engagement, creative research ideas, and multi-method and multi-disciplinary approaches, in addition to technical skill development and automated learning analytics methods.

Supporting access to different kinds of data

Along with curiosities about the 'bigger picture,' early career education researchers in our study frequently noted the value of combining learning analytics with other methods. Nearly all participants noted an interest and curiosity in experimenting with new approaches to answer complex research questions. By combining data from different methods, including qualitative methods, participants understood they could develop a more complete picture of learning than by only collecting log file data, which was in line with others in the field (Chatti et al., 2012; Suthers & Verbert, 2013). In this regard, combining multiple methods can help bridge the gap between generalisation and personalisation of visualisations and feedback to be used by students and teachers. However, there was a recognition of challenges in accessing and incorporating different types of data and the expense of using new technologies with large groups of students. Therefore, more support is needed to ensure that early career researchers have resources (time, access, training, etc.) to incorporate diverse methods into their research.

Developing and fostering ethical and theoretical engagements

To effectively incorporate learning analytics into their own research, early career education researchers need to be well-versed in current debates around ethics (Drachsler & Greller, 2016; Prinsloo & Slade, 2017b; Slade & Prinsloo, 2013). The early career researchers in this workshop appeared to be aware of these issues, but often were unsure about how to address them. Participants often felt they were small cogs in large institutional machines, with little power to influence or contribute to debates around ethics. Although the wider field is attempting to address these issues through a growing number of ethical frameworks and privacy guidelines (Gašević et al., 2016), early career researchers had major questions about how to collect data unbiasedly, while still capturing complex learning processes. Therefore, an explicit engagement with developing early career education researchers' knowledge of ethical frameworks and theories in learning analytics is vital. At the same time, the knowledges and experiences of early career researchers are important voices for institutions to consider when developing ethical frameworks.

5.1 Limitations

Through an in-depth analysis of an EARLI JURE conference workshop, this study has provided a nuanced look at the thoughts and experiences of early career education researchers in regards to learning analytics. We believe our study is the first of its kind to engage with how early career education researchers perceive and use learning analytics, providing insight into how the next generation of researchers will move this emerging field forward. In doing so, however, we note several limitations. First, it is recognised there was a potential for sample bias in relation to who volunteered to attend the workshop. It is likely that those who attended the workshop were already interested in and using learning analytics and that more diverse views could be elicited from early career researchers who have opted *not* to engage with the field. Future research, therefore, should explore our research questions with a broader population of early career education researchers in different contexts to identify barriers experienced by those with little to no experience or knowledge of learning analytics. Second, we note that the audio recordings for small group discussions were made in a busy, noisy room and, as such, were not always complete. Some participants' statements may have been omitted or unusable due to the quality of the recording. At the same time, discussion topics varied across the six small groups and it was not possible to elicit thoughts on some topics or themes from all participants. However, all main points from the small-group discussions were summarised by the facilitator, and all data were analysed and cross-examined by the seven early career researchers.

6. Conclusion

In this study, early career education researchers were aware of the strengths and limitations of learning analytics research and were grappling with previously identified issues in the field, including ethical and privacy issues, institutional barriers, concerns over atheoretical approaches, and developing complex stories about learning processes. This shows an awareness of field issues, as well as a familiarity with the literature and current research. The findings of this study also revealed how early career education researchers are struggling to situate themselves within the field and gain the complex skills necessary to appropriately embed learning analytics approaches into their own practices. These issues are essential considerations for the EARLI community and wider field of educational research, as it is clear that more resources are needed to support and develop the valuable expertise needed for early career researchers to contribute to the growing field of learning analytics.

Keypoints

- Early career researchers in this study reflected positively on using learning analytics in their research to support and understand students' learning processes
- Early career researchers recognised a variety of benefits and challenges of using learning analytics approaches in their research, which were frequently in line with the theorisation of experts in the field
- Most early career researchers favoured mixed methods approaches by combining learning analytics data with other quantitative and qualitative methods
- Early career researchers debated the role of theory in learning analytics and, in particular, whether data should support theory or theory should support data
- Common barriers to using learning analytics for early career researchers included access to data, technical skills in processing and analysing data, and contextualising findings to students' lives and experiences

References

Agudo-Peregrina, Á. F., Iglesias-Pradas, S., Conde-González, M. Á., & Hernández-García, Á. (2014). Can we predict success from log data in VLEs? Classification of interactions for learning analytics and their relation with performance in VLE-supported F2F and online learning. *Computers in Human Behavior, 31*, 542-550. doi: 10.1016/j.chb.2013.05.031

boyd, d., & Crawford, K. (2012). Critical questions for big data. *Information, Communication & Society, 15*(5), 662-679. doi: 10.1080/1369118X.2012.678878

Buckingham Shum, S., & Deakin Crick, R. (2012). *Learning dispositions and transferable competencies: Pedagogy, modelling and learning analytics*. Paper presented at the Learning Analytics and Knowledge (LAK), Vancouver, Canada.

Buckingham Shum, S., Hawksey, M., Baker, R., Jeffery, N., Behrens, J. T., & Pea, R. (2013). *Educational data scientists: A scarce breed*. Paper presented at the Learning Analytics and Knowledge (LAK), Leuven, Belgium.

Charleer, S., Klerkx, J., Duval, E., De Laet, T., & Verbert, K. (2016). *Creating effective learning analytics dashboards: Lessons learnt*. Paper presented at the European Conference on Technology Enhanced Learning (EC-TEL) Lyon, France.

Chatti, M. A., Dyckhoff, A. L., Schroeder, U., & Thüs, H. (2012). A reference model for learning analytics. *International Journal of Technology Enhanced Learning, 4*(5-6), 318-331. doi: 10.1504/ijtel.2012.051815

Creswell, J. W., & Miller, D. L. (2000). Determining validity in qualitative inquiry. *Theory Into Practice, 39*(3), 124-130. doi: 10.1207/s15430421tip3903_2

Dalziel, J. (2016). *Learning design: Conceptualizing a framework for teaching and learning online*. London, UK: Routledge.

Dawson, S., Gasevic, D., Siemens, G., & Joksimovic, S. (2014). *Current state and future trends: a citation network analysis of the learning analytics field*. Paper presented at the Proceedings of the Fourth International Conference on Learning Analytics And Knowledge, Indianapolis, Indiana, USA.

Dawson, S., & Siemens, G. (2014). Analytics to literacies: The development of a learning analytics framework for multiliteracies assessment. *International Review of Research in Open and Distributed Learning, 15*(4). doi: 10.19173/irrodl.v15i4.1878

Drachsler, H., & Greller, W. (2016). *Privacy and analytics: it's a DELICATE issue a checklist for trusted learning analytics*. Paper presented at the Learning Analytics and Knowledge (LAK), Edinburgh, United Kingdom.

Ferguson, R., Brasher, A., Cooper, A., Hillaire, G., Mittelmeier, J., Rienties, B., . . . Vuorikari, R. (2016). Research evidence of the use of learning analytics: Implications for education policy. In R. Vuorikari & J. Castano-Munoz (Eds.), *A European Framework for Action on Learning Analytics*. Luxembourg: Joint Research Centre Science for Policy Report.

Ferguson, R., Clow, D., Macfadyen, L., Essa, A., Dawson, S., & Alexander, S. (2014). *Setting learning analytics in context: Overcoming the barriers to large-scale adoption*. Paper presented at the Learning Analytics and Knowledge (LAK), Indianapolis, Indiana, USA.

García, R. M. C., Pardo, A., Kloos, C. D., Niemann, K., Scheffel, M., & Wolpers, M. (2012). Peeking into the black box: Visualising learning activities. *International Journal of Technology Enhanced Learning, 4*(1-2), 99-120. doi: 10.1504/ijtel.2012.048313

Gašević, D., Dawson, S., & Jovanovic, J. (2016). Ethics and privacy as enablers of learning analytics. *Journal of Learning Analytics, 3*(1), 4. doi: 10.18608/jla.2016.31.1

Gašević, D., Dawson, S., & Siemens, G. (2015). Let's not forget: Learning analytics are about learning. *TechTrends, 59*(1), 64-71. doi: 10.1007/s11528-014-0822-x

Ifenthaler, D., & Widanapathirana, C. (2014). Development and validation of a learning analytics framework: Two case studies using support vector machines. *Technology, Knowledge and Learning, 19*(1), 221-240. doi: 10.1007/s10758-014-9226-4

Kirschner, P. (2016). *Learning analytics: Utopia or dystopia* Paper presented at the Learning Analytics and Knowledge (LAK) Keynote. http://lak16.solaresearch.org/?events=keynote-slides-available

Kirschner, P., & Erkens, G. (2013). Toward a framework for CSCL research. *Educational Psychologist, 48*(1), 1-8. doi: 10.1080/00461520.2012.750227

Kovanovic, V., Gašević, D., Dawson, S., Joksimovic, S., & Baker, R. (2016). Does Time-on-task Estimation Matter? Implications on Validity of Learning Analytics Findings. *2016, 2*(3), 81-110. doi: 10.18608/jla.2015.23.6

Kovanovic, V., Gašević, D., Dawson, S., Joksimović, S., Baker, R. S., & Hatala, M. (2015). *Penetrating the black box of time-on-task estimation.* Paper presented at the Learning Analytics and Knowledge (LAK), Poughkeepsie, New York, USA.

LAK. (2011). Call for Papers: 1st Annual Conference on Learning Analytics and Knowledge.

Lang, C., Siemens, G., Wise, A. F., & Gašević, D. (2017). *Handbook of learning analytics*: Society for Learning Analytics Research.

Lichtman, M. (2013). *Qualitative research in education: A user's guide* (3 ed.). Thousand Oaks, CA: SAGE.

Lockyer, L., Heathcote, E., & Dawson, S. (2013). Informing pedagogical action: Aligning learning analytics with learning design. *American Behavioral Scientist, 57*(10), 1439-1459. doi: 10.1177/0002764213479367

Malmberg, J., Järvelä, S., & Järvenoja, H. (2017). Capturing temporal and sequential patterns of self-, co-, and socially shared regulation in the context of collaborative learning. *Contemporary Educational Psychology, 49*, 160-174. doi: 10.1016/j.cedpsych.2017.01.009

Merceron, A., Blikstein, P., & Siemens, G. (2016). Learning analytics: From big Data to meaningful data. *Journal of Learning Analytics, 2*(3), 5. doi: 10.18608/jla.2015.23.2

Mirriahi, N., & Vigentini, L. (2017). Analytics of learner video use. In C. Lang, G. Siemens, A. F. Wise, & D. Gasevic (Eds.), *Handbook of learning analytics* (pp. 49-57): Society for Learning Analytics Research.

Nguyen, Q., Rienties, B., Toetenel, L., Ferguson, R., & Whitelock, D. (2017). Examining the designs of computer-based assessment and its impact on student engagement, satisfaction, and pass rates. *Computers in Human Behavior, 76*, 703-714. doi: 10.1016/j.chb.2017.03.028

Papamitsiou, Z., & Economides, A. (2016). Learning analytics for smart learning environments: A meta-analysis of empirical research results from 2009 to 2015. In J. M. Spector, B. B. Lockee, & D. M. Childress (Eds.), *Learning, design, and technology: An international compendium of theory, research, practice, and policy* (pp. 1-23). London, UK: Springer.

Persico, D., & Pozzi, F. (2015). Informing learning design with learning analytics to improve teacher inquiry. *British Journal of Educational Technology, 46*(2), 230-248. doi: 10.1111/bjet.12207

Prinsloo, P., & Slade, S. (2017a). *An elephant in the learning analytics room: The obligation to act*. Paper presented at the Learning Analytics & Knowledge (LAK), Vancouver, Canada.

Prinsloo, P., & Slade, S. (2017b). Ethics and learning analytics: Charting the (un)charted. In C. Lang, G. Siemens, A. F. Wise, & D. Gasevic (Eds.), *Handbook of learning analytics* (pp. 49-57): Society for Learning Analytics Research.

Rienties, B., & Toetenel, L. (2016). The impact of learning design on student behaviour, satisfaction and performance: A cross-institutional comparison across 151 modules. *Computers in Human Behavior, 60*, 333-341. doi: 10.1016/j.chb.2016.02.074

Ruppert, E., Harvey, P., Lury, C., Mackenzie, A., McNally, R., Baker, S. A., & Lewis, C. (2015). *Socialising big data: from concept to practice*. Working Paper Series (No. 138): Centre for Research on Socio-Cultural Change.

Scholes, V. (2016). The ethics of using learning analytics to categorize students on risk. *Educational Technology Research and Development*, *64*(5), 939-955. doi: 10.1007/s11423-016-9458-1

Selwyn, N. (2015). Data entry: Towards the critical study of digital data and education. *Learning, Media and Technology*, *40*(1), 64-82. doi: 10.1080/17439884.2014.921628

Selwyn, N. (2016). *Education and technology: Key issues and debates*. London: Bloomsbury Publishing.

Sharples, M., de Roock, R., Ferguson, R., Gaved, M., Herodotou, C., Koh, E., . . . Wong, L. H. (2016). Innovating Pedagogy 2016: Open University Innovation Report 5. Milton Keynes: The Open University.

Siemens, G. (2013). Learning analytics: The emergence of a discipline. *American Behavioral Scientist, 57*(10), 1380-1400. doi: 10.1177/0002764213498851

Slade, S., & Prinsloo, P. (2013). Learning analytics: Ethical issues and dilemmas. *American Behavioral Scientist*, *57*(10), 1510-1529. doi: 10.1177/0002764213479366

Suthers, D., & Verbert, K. (2013). *Learning analytics as a "middle space"*. Paper presented at the Proceedings of the Third International Conference on Learning Analytics and Knowledge, Leuven, Belgium.

Tempelaar, D. T., Rienties, B., & Giesbers, B. (2015). In search for the most informative data for feedback generation: Learning Analytics in a data-rich context. *Computers in Human Behavior, 47*, 157-167. doi: 10.1016/j.chb.2014.05.038

Tempelaar, D. T., Rienties, B., Mittelmeier, J., & Nguyen, Q. (2017). Student profiling in a dispositional learning analytics application using formative assessment. *Computers in Human Behavior, 78*, 408-420. doi: 10.1016/j.chb.2017.08.010

van Leeuwen, A., Janssen, J., Erkens, G., & Brekelmans, M. (2015). Teacher regulation of cognitive activities during student collaboration: Effects of learning analytics. *Computers & Education, 90*, 80-94. doi: 10.1016/j.compedu.2015.09.006

Winne, P. H. (2017). Leveraging big data to help each learner upgrade learning and accelerate learning science. *Teachers College Record, 119*(13), 1-24.

Appendix 1: Thematic analysis codes

Theme	Sub-theme	Example quote
LA can be used to afford student success (focus on SRL)	Combining learning data and learner data provides opportunities to support student success	'The project we presented at LAK was looking at what can we tell students who are already doing okay but want to do better? And we have found it tends to be completely different things. '
	Provides feedback into the learning processes to capture when students adjust their learning approaches	'But that is what we want to look at is how students can look at their challenges, see patterns, and see what strategies they may change or how they seek help once they see the patterns in themselves or others.'
	Focus on non-demographic data (i.e. behaviour) so the focus is on what students can change about themselves as they learn	'There is a lot of use of demographic data in LA and I like what Phil Winne said in the Handbook of Learning Analytics was that he thinks the focus should be on things learners can change about themselves.'
	'Selling' learning analytics as beneficial to students and not just advisors and teachers	'I don't mind selling it because with my background, I know how it can benefit students. At our institution, when they first were talking about LA, they were focusing on advisors, not students or teachers.'
	Focusing on preventing failure in students at risk, but also promoting success in students with passing/good/excellent grades	'I think that is what we were seeing is preventing students from failing, but the project we presented at LAK was looking at what can we tell students who are already doing okay but want to do better?'
	Danger of LA doing the regulating for students rather than students regulating their own learning	'Also, we talk a lot about "who is doing the regulating?" And how do you scaffold it, so it isn't given the same level of prompts all the time?'
Learning analytics is on the cutting edge of educational research	Learning analytics techniques are unfamiliar to some early career researchers, but they believe they could be useful	'I have a question actually about the definition of learning analytics. Does it have to be big data, or can it be any kind of numeric data?'
	Trace files data can be challenging to interpret and analyse	'Well I think that the log files we can get from the LMS is easily accessible, but in the literature it's a bit unclear about what that data actually mean. Sometimes it's not the amount of time they spent but the sequence of time they spent on something…'
	Challenging to establish (or find) the expertise needed	'I'm now working as a research associate on the learning analytics project at my institution and I am the only person along with my collaborator. We are the experts on learning analytics at our university.' 'For right now, it's just me and I have to find everyone. So, I'm building it from the ground up and I have to find people at my institution with these skills.'
	Challenging to access the technologies needed to implement	'So, the next part is that we don't have any visualizations or dashboards. That is one of our problems is how to implement that'
Multiple Methods	Some early career researchers already combined learning analytics with other methods	'The analytics part, I hope is going to come with social network analysis. And then I'm also planning to supplement it with interviews and things like that.'
	Combining learning analytics with other methods provides a stronger evidence base and can reduce bias	(interaction) P1: 'They still talk face-to-face, and they still have informal communication networks. How do we bring that into learning analytics research?' P2: 'Qualitative?' [spoken hesitantly]
	Using multiple methods can provide the 'how-and-why' (the intent)	'But it is all really flat at the moment. I ran in a lot of trouble of, like, getting a lot of data but not really feel how to interpret it. Because you only see the patterns, but you don't know why or how … I'm really think about mixing it with the log data analysis but then also interviewing them … You usually need somebody's explanation about why they did it what they did.'
	Can be challenging to get access to data needed to use multiple methods	'Me, probably the biggest struggle is access, because I am working, maybe this is not relevant to other, but I'm working with companies. So being able to negotiate times for interviews versus being able to access potentially sensitive information. It takes a lot of times just to negotiate access to certain types of resources and, as you say, time is limited'

Theory	Combining learning data and learner data provides opportunities for learning analytics to draw on learning theory	'That's what has attracted me to learning analytics because there is an aim to include more educational theorists and it's a solid group of people from all disciplines.'
	Early career researchers debate whether learning analytics should be theory driven or data driven.	'…That sometimes works, but the theory often seduces you to see something in the data which might not entirely be there. It's still quite tricky…'
	Theory can provide the 'why'	'I have noticed in the learning analytics literature that often … it doesn't have the theory, like why certain variables are chosen and that kind of why piece …You need to, like, think a little bit more about why those things might cause those relationships to appear'
Contextualization	Context of where/how the data was collected is important; finding may not translate between contexts	'Yeah, and the contextualisation of your data is quite important' 'Only if the context is described, [mumbled] and it matches, then it might be able to follow'
	You can't always get access to the all the data because of the context in which it was collected	'Well I think part of it is like in an ideal world, data is collected from the LMS so a lot of times in other platforms that people are using, we don't have access to all of those'
	Data needs to be contextualized	'Sometimes it's not the amount of time they spent but the sequence of time they spent on something, or you have to know the time when they have an exam, so you need to take into account the context of the course, so it is not just [words] or something.'
Ethics	Learning analytics carries a risk that students will be stereotyped, this can be avoided but may be a necessary evil. And, is categorizing stereotyping?	'It's about the argumentation of why you focus on the women rather than the men, maybe. For example, if the beneficial effects are larger for women than for men, then I would say, yeah, maybe. But, then again....'
	It may not be ethical to track all variables	'Yeah, I don't like it [using gender as a variable]. Why should we have to do that all the time?'
	Learning analytics affords collection of data on everyone: Affords less bias in interventions	'I think that is what we were seeing is preventing students from failing, but the project we presented at LAK was looking at what can we tell students who are already doing okay but want to do better?'
	Protection of privacy and individual rights	'We really have big problems I think with ethics because we can't actually not look at a lot of data. This is because we have to protect students' rights.'
	Student Consent	'And of course, the students always have the right to not give away their data. So, we always have to ask them to do it.'
	Institutions are risk adverse	'I think one of the big problems in our institution is that the institution really blocks everything in that direction.' 'We can't invest so much money, so why opening this up to someone who wants to do research with this when we have the danger of being sued.'
	Community concerns around data use	'Yes, there are some concerns from the community about what kinds of data are you collecting? What/why are you using it for? and very importantly what are the implications if your personal data is being collected, even if it is for something for a very noble cause to teaching and learning. This is just democratic.'
	Ethics surrounding control groups	'…normally I should some kind of really controlled group. How do I ethically handle with that I don't give them the best opportunity to learn?'
	Concerns regarding unintended consequences	'Personally, I think one concern of the ethical implications would be that of unintended consequences.' 'Some of these issues are like a double-edged sword. This is powerful but it could be used destructively.'
	Stakeholders are not prepared for the power of learning analytics	'A possible consequence of that because things are more visible, so maybe teachers who then become afraid to try to innovative, particularly because doing so in a challenging academic field cost time, and this failure sometimes cannot be acceptable in some institutions.'